

CRAG 7-58
2 April 1958

CENTRAL INTELLIGENCE AGENCY

CENTRAL REFERENCE ADVISORY GROUP

Selective Indexing of Intelligence Publications by OCR

OCR receives annually approximately 400,000 different intelligence reports, periodicals and monographic studies from the IAC Agencies, from IAC contractors, from other government or private sources and from certain foreign governments.

As each series has appeared on the scene for the first time or as documents with unusual subject content have been identified in the daily receipts, OCR has had to adjust its indexing policy to accommodate some and to exclude others depending on the apparent intelligence significance of the material and its acceptability to the OCR processing system. Many of the decisions have been discussed with one or more using offices, however, no comprehensive set of selection criteria has been evolved and publicized to insure that all who use the system are properly aware of these limitations.

The attached paper sets forth some of the principal considerations bearing on selective indexing of intelligence publications and proposes that CRAG establish a Working Group to review and confirm existing criteria and to recommend measures for improvement.

Paul A. Borel
Paul A. Borel
Assistant Director
Central Reference

Attachment:

Memorandum for AD/CR, dated 29 March 1958,

Subject: CRAG Review of OCR's Selective Indexing of Intelligence Publications

31 March 1958

MEMORANDUM FOR: AD/CR

SUBJECT : CRAG Review of OCR's Selective Indexing of Intelligence Publications

I. PROBLEM

To clarify policy on selective indexing of intelligence literature into the OCR Intellofax system.

II. BACKGROUND

OCR receives annually approximately 400,000 different intelligence reports, periodicals and monographic studies from the IAC Agencies, from IAC contractors, from other government or private sources and from certain foreign governments.

Two distinct processing operations are applied by OCR to this flow, (1) copies of each report are disseminated to interested offices according to their standing subject requirements; (2) one copy of each report is given some or all of the following processing to provide physical control and subject retrievability within the Intellofax system:

log	-	source cards, Kardex, dissemination record
index	-	IBM card entries in Intellofax file or entries in monthly, printed IPI
film	-	aperture cards or microfilm reels
store	-	non-filmed items stored on library shelves

C-O-N-F-I-D-E-N-T-I-A-L

- 2 -

III. DISCUSSION

A. The Indexing Decision

The indexing policies covering the processing and use of a very large document flow are obviously of importance to analysts in planning their basic methods of operation, e.g., nature and size of local files, and in conducting research on the given project. They are important to processors in establishing clear procedures, achieving consistency of treatment and in producing end results which satisfy users.

The questions of when to index a document and how to do it need to be examined from at least three points of view:

1. Is there a characteristic physical format which modifies the indexing approach?
2. Is there a characteristic subject scope which modifies the indexing approach?
3. What is the apparent value of the series or document to users of the system?

C-O-N-F-I-D-E-N-T-I-A-L

- 3 -

1. Representative types of documents having characteristic physical format which modifies the indexing approach:

Information Reports

(also physically unique and unstandardized as to source agency and series)

Enclosures

(books, periodicals, correspondence ---- artifacts)

Translations

Interrogations

Memoranda of Conversation

Cables

Periodicals, Press Summaries

Indexes, Accessions Lists, Bibliographies

Abstracts (book or card form), Information Series on cards

Statistical Tables, Target Sheets

Photographs, Microfilm

Maps, Charts

C-O-N-F-I-D-E-N-T-I-A-L

2. Representative categories of documents having characteristic subject scope which modifies the indexing approach:

Collection Requirements

Responses to Collection Requirements

Other Field Reporting

Travel Reports

Agricultural Attache Reports for the Department of Agriculture

Commercial Attache Reports for the Department of Commerce

Intelligence Periodicals and Periodic Reporting Within Information

Report Series

Administrative Communications - via separate series or in Information

Report Series

re: Quarters

Personnel

Funds

Plans

Transmittals

Status and Final Reports on Contractual Projects

Evaluations

Monographic Studies

Estimates

Sensitive activities requiring unusual limitations on access to information

C-O-N-F-I-D-E-N-T-I-A-L

3. Measurements of value of intelligence publications to users:

General Categories

a. Documents of no value for intelligence purposes:

Applies for example to communications between an agency's headquarters and its field establishments transmitted in regular information report or despatch channels - administrative in character, e.g., promotions of personnel.

b. Documents of interest to specialized users who obtain copies of all items received and operate local information storage and retrieval systems - e.g., "Strategic Trade Files" (in RR/E/ST) - includes, for example, "Trader-name Subdivision, a Comprehensive File of Extracts on Firms and Individuals Unique in Washington."

c. Documents containing specialized data but of general interest:

May be suited to indexing in the general (Intellofax) system but may be better suited to specialized treatment as in the OCR Biographic, Industrial, and Graphics Registers.

Some information, e.g. statistical data, order of battle, photographs or FBIS briefs may be suited only to a single Register system or to some external system, e.g. newspaper index, not presently maintained in government.

d. Documents of general interest and suited to indexing in the general (Intellofax) system.

C-O-N-F-I-D-E-N-T-I-A-L

- 6 -

Effect of Time on the Scale of Values

- e. Documents of exceptionally brief and transitory interest - e.g. daily status reports generally superseded by succeeding reports; minutes of committees; cables after an initial response is replaced by a despatch dealing with the problem in detail. May combine with any one of categories a - d above.
- f. Documents of no current interest but potentially significant in the future: Information reporting from an apparently stable area which will receive extensive re-examination in the event of political disturbance. May combine with any one of categories, a - d above.
- g. Documents of apparent general and continuing interest but subject to obsolescence due to change in the climate of community activities, for example, examination of a five year file is likely to yield decisions to retain all data on communism but to retire much of the data on plant, products, research, etc.

Economic Considerations

- h. Documents of special or general interest, suited to the general (Intellofax) system but received in very large volume or requiring extra processing and therefore exceeding the processing capacity of the system. Some recombination of factors in the system will be required - e.g. increase productivity or expand staff, or, application of priorities to eliminate this or some other category of material from the system, or, establishment of some abbreviated pattern for all or

- 7 -

Given receipt of a new document series or of documents within an established series dealing with new subjects, the OCR system must generate the necessary processing decisions in approximately the following manner:

- a. The series or document must be identified as a problem requiring special treatment. This is accomplished by instructions to head-of-the-line screeners.
- b. Series or documents must be categorized by format or subject scope as suited to specialized processing only, e.g. Map Library, Graphics Register, or to multiple indexing in the specialized and the general systems.
- c. Series or documents categorized as suited to general indexing must be further segregated, i.e.:
 1. As of marginal value or manifest value to known users of the system.
 2. As to compatibility of the subject content with the subject classification scheme (ISC); new subjects will require expansion or amendment of the scheme.
 3. Suitability of the document to system processes, e.g. legibility, photographability, special handling necessitated by unusual security classifications.

C-O-N-F-I-D-E-N-T-I-A-L

- 8 -

B. Resume of the Existing Situation

OCR disposes currently of hundreds of series of documents and thousands of individual documents based on decisions established along the procedural lines described above. The results in terms of indexing coverage in 1957 were as follows:

documents indexed	234,116	
documents not indexed (considered but rejected)	73,336	
periodicals and miscellaneous IAC series	69,116	<u>1/</u>
non-IAC series (UN, ICA, etc.)	<u>20,967</u>	<u>1/</u>
TOTAL	397,535	
cables (not indexed)	<u>157,773</u>	
GRAND TOTAL	555,308	

1/ Indexed selectively in the Intelligence Publications Index (IPI) or in the Intellofax system or in both.

C-O-N-F-I-D-E-N-T-I-A-L

- 9 -

C. Validation of the System

Recent studies of the OCR document processing system have urged that the decision not to index is of vital concern to the entire intelligence community and that OCR should arrange for community validation of all such decisions.

The preceding discussion outlines the principal elements of the decision making process as a basis for consultations with using offices concerning existing practices, improvements, and measures to be taken to publicize the system.

It points out that no single indexing system is adequate to the control of all of the series of documents received in OCR and also suggests that the indexing of some series cannot be justified.

It is important to note that the decision not to index does not mean denial of the document to the user nor elimination of the document from the OCR system. The typical document is received in OCR in 15 copies and all, including the processing copy if filmed, are disseminated in the Agency and may eventually be absorbed into a few or many working files. Secondly, OCR provides physical storage and retrieval service for all intelligence documents whether indexed or not. Non-IAC series may be screened to select items of intelligence interest for processing or they may be discontinued in toto if there is no known demand for the information, however, procurement will be resumed upon request.

The relative and subjective nature of user requirements for a series or a document is obviously the crux of the problem. An orderly administration of a system as large as OCR's would appear to be out of the question if unanimous approval of each indexing decision were sought.

While the responsibility for final decision must be OCR's, the scale of

Approved For Release 2000/05/23 : CIA-RDP67-00896R000100160002-5

- 10 -

values and the method of arriving at each decision are of equal importance to processors and to users and ought to be fully debated and publicized.

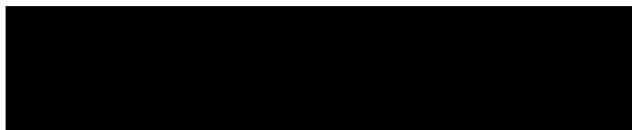
IV CONCLUSIONS

- A. The physical variations in publications received in OCR preclude indexing under a single, uniform system.
- B. Variations in subject scope of publications preclude indexing under a single, uniform system.
- C. Some categories of publications are compatible with one or more of the principal OCR indexing systems but are of no current interest to users. Limited OCR resources, at present, do not permit full-scale indexing of all marginal categories.
- D. While OCR indexing policies are selective and non-universal in character, they are not fully codified nor publicized. A general review of existing practice might alter present indexing priorities and generate more efficient methods for indoctrinating intelligence personnel in the use of the system.

V RECOMMENDATION

That CRAG establish a Working Group made up of representatives from appropriate CIA offices to review and confirm OCR indexing policy or to recommend measures for its improvement.

25X1A9a



Chief, Document Division

C-O-N-F-I-D-E-N-T-I-A-L